

# HIGH DEGREES IN RANDOM RECURSIVE TREES

LOUIGI ADDARIO-BERRY AND LAURA ESLAVA

**ABSTRACT.** For  $n \geq 1$ , let  $T_n$  be a random recursive tree (RRT) on the vertex set  $[n] = \{1, \dots, n\}$ . Let  $\deg_{T_n}(v)$  be the degree of vertex  $v$  in  $T_n$ , that is, the number of children of  $v$  in  $T_n$ . Devroye and Lu [6] showed that the maximum degree  $\Delta_n$  of  $T_n$  satisfies  $\Delta_n / \lfloor \log_2 n \rfloor \rightarrow 1$  almost surely; Goh and Schmutz [7] showed distributional convergence of  $\Delta_n - \lfloor \log_2 n \rfloor$  along suitable subsequences. In this work we show how a version of Kingman's coalescent can be used to access much finer properties of the degree distribution in  $T_n$ .

For any  $i \in \mathbb{Z}$ , let  $X_i^{(n)} = |\{v \in [n] : \deg_{T_n}(v) = \lfloor \log n \rfloor + i\}|$ . Also, let  $\mathcal{P}$  be a Poisson point process on  $\mathbb{R}$  with rate function  $\lambda(x) = 2^{-x} \cdot \ln 2$ . We show that, up to lattice effects, the vectors  $(X_i^{(n)}, i \in \mathbb{Z})$  converge weakly in distribution to  $(\mathcal{P}[i, i+1), i \in \mathbb{Z})$ . We also prove asymptotic normality of  $X_i^{(n)}$  when  $i = i(n) \rightarrow -\infty$  slowly, and obtain precise asymptotics for  $\mathbf{P}(\Delta_n - \log_2 n > i)$  when  $i(n) \rightarrow \infty$  and  $i(n)/\log n$  is not too large. Our results recover and extends the previous distributional convergence results on maximal and near-maximal degrees in random recursive trees.

## 1. STATEMENT OF RESULTS

The process of random recursive trees  $(T_n, n \geq 1)$  is defined as follows.  $T_1$  has a single node with label 1, which is its root. The tree  $T_{n+1}$  is obtained from  $T_n$  by directing an edge from a new vertex  $n+1$  to  $v \in [n]$ ; the choice of  $v$  is uniformly random and independent for each  $n \in \mathbb{N}$ . We call  $T_n$  a random recursive tree (RRT) of size  $n$ .

As a consequence of the construction, vertex-labels in  $T_n$  increase along root-to-leaf paths. Rooted labelled trees with such property are called *increasing trees*. It is not difficult to see that, in fact,  $T_n$  is uniformly chosen among the set  $\mathcal{T}_n$  of increasing trees with vertex set  $[n]$ .

We write  $\deg_{T_n}(v)$  to denote the number of children of  $v$  in  $T_n$ . The degree distribution of  $T_n$  is encoded by the variables  $Z_i^{(n)} = |\{v \in [n] : \deg_{T_n}(v) = i\}|$ , for  $i \geq 0$ . In fact, the study of RRT's started with a paper by Na and Rapoport [13] in which they obtained, for any *fixed*  $i \geq 0$ , the convergence  $\mathbb{E}(Z_i^{(n)})/n \rightarrow 2^{-i-1}$  as  $n \rightarrow \infty$ ; this result was extended to convergence in probability by Meir and Moon in [12]. Mahmoud and Smythe [11] derived the asymptotic joint normality of  $Z_i^{(n)}$  for  $i \in \{0, 1, 2\}$ ; and finally, Janson [8] extended the joint normality to  $Z_i^{(n)}$  for  $i \geq 0$  and gave explicit formulae for the covariance matrix.

The above results concern typical degrees; the focus in this work is large degree vertices, and in particular the maximum degree in  $T_n$ , which we denote  $\Delta_n = \max_{v \in [n]} \deg_{T_n}(v)$ . For the rest of the paper we write  $\log$  to denote logarithms with base 2, and  $\ln$  to denote natural logarithms. For  $n \in \mathbb{N}$  let  $\varepsilon_n = \log n - \lfloor \log n \rfloor$ .

A heuristic to find the order of  $\Delta_n$  is that, if  $\mathbb{E}(Z_i^{(n)}) \approx n2^{-i-1}$  were to hold for all  $i$ , as it does when  $i$  is fixed, then we would have  $\mathbb{E}(Z_{\lfloor \log n \rfloor}^{(n)}) \approx n2^{-\lfloor \log n \rfloor - 1} = 2^{-1+\varepsilon_n}$ .

---

Date: July 4, 2016.

2010 *Mathematics Subject Classification.* 60C05, 05C80.

This heuristic suggests that  $\Delta_n$  is of order  $\log n$ . This is indeed the case: Szymanski [15] proved that  $\mathbf{E}[\Delta_n]/\log n \rightarrow 1$  as  $n \rightarrow \infty$ , and Devroye and Lu [6] later established that  $\Delta_n/\log n \rightarrow 1$  a.s.. Finally, Goh and Schmutz [7] showed that  $\Delta_n - \lfloor \log n \rfloor$  converges in distribution along suitable subsequences, and identified the possible limiting laws.

Since we focus on maximal degrees, it is useful to let

$$X_i^{(n)} = Z_{i+\lfloor \log n \rfloor}^{(n)} = |\{v \in [n] : \deg_{T_n}(v) = \lfloor \log n \rfloor + i\}|,$$

for  $n \in \mathbb{N}$  and  $i \geq -\lfloor \log n \rfloor$ . The following is a simplified version of one of our main results.

**Theorem 1.1.** *Fix  $\varepsilon \in [0, 1]$ . Let  $(n_l)_{l \geq 1}$  be an increasing sequence of integers satisfying  $\varepsilon_{n_l} \rightarrow \varepsilon$  as  $l \rightarrow \infty$ . Then, as  $l \rightarrow \infty$*

$$(X_i^{(n_l)}, i \in \mathbb{Z}) \xrightarrow{d} (P_i^\varepsilon, i \in \mathbb{Z})$$

*jointly for all  $i \in \mathbb{Z}$  where the  $P_i^\varepsilon$  are independent Poisson r.v.'s with mean  $2^{-i-1+\varepsilon}$ .*

The random variables  $X_i^{(n)}$  do not converge in distribution as  $n \rightarrow \infty$  without taking subsequences; this is essentially a lattice effect caused by the floor  $\lfloor \log n \rfloor$  in the definition of  $X_i^{(n)}$ .

Theorem 1.1 can be stated in terms of weak convergence of point processes (which is equivalent to convergence of finite dimensional distributions (FDD's); see Theorem 11.1.VII in [4]). In fact, we will also prove convergence (along subsequences) of

$$X_{\geq i}^{(n)} = \sum_{k \geq i} X_k^{(n)} = |\{v \in [n] : \deg_{T_n}(v) \geq \lfloor \log n \rfloor + i\}|.$$

This is useful as it yields information about  $\Delta_n$  which cannot be derived from Theorem 1.1. We formulate this result as a statement about convergence of point processes, and now provide the relevant definitions. Let  $\mathbb{Z}^* = \mathbb{Z} \cup \{\infty\}$ . Endow  $\mathbb{Z}^*$  with the metric defined by  $d(i, j) = |2^{-j} - 2^{-i}|$  and  $d(i, \infty) = 2^{-i}$  for  $i, j \in \mathbb{Z}$ . Let  $\mathcal{M}_{\mathbb{Z}^*}^\#$  be the space of boundedly finite measures of  $\mathbb{Z}^*$ .

Let  $\mathcal{P}$  be a Poisson point process on  $\mathbb{R}$  with rate function  $\lambda(x) = 2^{-x} \cdot \ln 2$ . For each  $\varepsilon \in [0, 1]$  let  $\mathcal{P}^\varepsilon$  be the point process on  $\mathbb{Z}^*$  given by

$$\mathcal{P}^\varepsilon = \sum_{x \in \mathcal{P}} \delta_{\lfloor x + \varepsilon \rfloor}.$$

Similarly, for all  $n \in \mathbb{N}$  let

$$\mathcal{P}^{(n)} = \sum_{v \in [n]} \delta_{\deg_{T_n}(v) - \lfloor \log n \rfloor}.$$

Then, for each  $i \in \mathbb{Z}$  we have that

$$\mathcal{P}^\varepsilon(\{i\}) := |\{x \in \mathcal{P} : \lfloor x + \varepsilon \rfloor = i\}| = |\{x \in \mathcal{P} : x \in [i - \varepsilon, i + 1 - \varepsilon)\}|$$

has distribution  $\text{Poi}(2^{-i-1+\varepsilon})$ ; also  $\mathcal{P}^{(n)}(\{i\}) = X_i^{(n)}$ . We abuse notation by writing, e.g.,  $\mathcal{P}^{(n)}(i) = \mathcal{P}^{(n)}(\{i\})$ .

It is clear that  $\mathcal{P}^{(n)}$  and  $\mathcal{P}^\varepsilon$  are elements of  $\mathcal{M}_{\mathbb{Z}^*}^\#$ . The advantage of working on the state space to  $\mathbb{Z}^*$  is that intervals  $[k, \infty]$  are compact. In particular, the convergence of FDD's of  $\mathcal{P}^{(n_l)}$  implies the convergence in distribution of  $X_{\geq i}^{(n_l)} = \mathcal{P}^{(n_l)}[i, \infty)$ .

**Theorem 1.2.** Fix  $\varepsilon \in [0, 1]$ . Let  $(n_l)_{l \geq 1}$  be an increasing sequence of integers satisfying  $\varepsilon_{n_l} \rightarrow \varepsilon$  as  $l \rightarrow \infty$ . Then in  $\mathcal{M}_{\mathbb{Z}^*}^\#$ ,  $\mathcal{P}^{(n_l)}$  converges weakly to  $\mathcal{P}^\varepsilon$  as  $l \rightarrow \infty$ . Equivalently, for any  $i < i' \in \mathbb{Z}$ , jointly as  $l \rightarrow \infty$

$$(X_i^{(n_l)}, \dots, X_{i'-1}^{(n_l)}, X_{\geq i'}^{(n_l)}) \xrightarrow{d} (\mathcal{P}^\varepsilon(i), \dots, \mathcal{P}^\varepsilon(i' - 1), \mathcal{P}^\varepsilon[i', \infty)).$$

Note that Theorem 1.1 follows from Theorem 1.2. We finish this section stating two additional results. The first is an extension of the main theorem from [7], that result being essentially the case  $i = O(1)$ .

**Theorem 1.3.** For any  $i = i(n)$  with  $i + \log n < 2 \ln n$  and  $\liminf_{n \rightarrow \infty} i(n) > -\infty$ ,

$$\mathbf{P}(\Delta_n \geq \lfloor \log n \rfloor + i) = (1 - \exp\{-2^{-i+\varepsilon_n}\})(1 + o(1)).$$

When  $i = O(1)$ , the assertion of Theorem 1.3 is a straight-forward consequence of Theorem 1.2. For the case that  $i(n) \rightarrow \infty$  we use estimates for the first and second moments of  $X_{\geq i}^{(n)}$ ; note that  $\{\Delta_n < \lfloor \log n \rfloor + i\} = \{X_{\geq i}^{(n)} = 0\}$ .

Finally, we also obtain the asymptotic normality for  $X_i^{(n)}$  when  $i$  tends to  $-\infty$  slowly enough.

**Theorem 1.4.** If  $i = i(n) \rightarrow -\infty$  and  $i = o(\ln n)$ , then as  $n \rightarrow \infty$

$$\frac{X_i^{(n)} - 2^{-i-1+\varepsilon_n}}{\sqrt{2^{-i-1+\varepsilon_n}}} \xrightarrow{d} N(0, 1).$$

**Remark 1.5.** Up to lattice effects, Theorems 1.2 and 1.4 extend the range of  $i = i(n)$  for which the heuristic that  $Z_i^{(n)} \approx n2^{-i-1}$  holds.

A key novelty of our approach is that for each  $n$  we use *Kingman's coalescent* to generate a tree  $T^{(n)}$  whose vertex degrees  $\{\deg_{T^{(n)}}(v)\}_{v \in [n]}$  are exchangeable but otherwise have the same law as degrees in  $T_n$ . (See [2], Chapter 2 for a description of Kingman's coalescent, and [1], Section 2.2 for a description of the connection with random recursive trees which we exploit in this paper.) By this we mean that if  $\sigma : [n] \rightarrow [n]$  is a uniformly random permutation then the following distributional identity holds:

$$(1) \quad (\deg_{T^{(n)}}(v), v \in [n]) \stackrel{d}{=} (\deg_{T_n}(\sigma(v)), v \in [n]).$$

We describe the trees  $T^{(n)}$ ,  $n \in \mathbb{N}$  in Section 3.

An essentially equivalent construction was used by Devroye [5] to study union-find trees. In [14], Pittel related the results of [5] on union-find trees to the height of RRT's. It is worth mentioning that both Kingman's coalescent and the union-find trees can be equivalently represented as binary trees or, as we will see in Section 3, as RRT's. Aside from the works [5] and [14], it seems that the use of Kingman's coalescent or of union-find trees to study RRT's is rare. However, it turns out to provide just the right perspective for studying high degree vertices.

## 2. OUTLINE

In this section we sketch the approach used in the paper. The proofs of the theorems relay on the computation of the moments of the FDD's of  $\mathcal{P}^{(n)}$ ; these estimates are given in Proposition 2.1. In particular, the proofs of Theorems 1.2 and 1.4 use the method of moments (e.g., see [9] Section 6.1, and [3] Section 1.5).

Any FDD of  $\mathcal{P}^{(n)}$  can be recovered from suitable marginals of the joint distribution of  $(X_i^{(n_l)}, \dots, X_{i'-1}^{(n_l)}, X_{\geq i'}^{(n_l)})$  for some  $i < i' \in \mathbb{Z}$ . For simplicity, we focus for the moment

on collections of variables  $X_i^{(n)}, \dots, X_{i'}^{(n)}$  for  $i \leq i'$ . For  $r \in \mathbb{R}$  and  $a \in \mathbb{N}$  write  $(r)_a = r(r-1)\dots(r-a+1)$ , also let  $(r)_0 = 1$ . We will prove that for any non-negative integers  $a_i, \dots, a_{i'}$ , as  $n \rightarrow \infty$ , we have

$$(2) \quad \mathbf{E} \left[ \prod_{i \leq k \leq i'} (X_k^{(n)})_{a_k} \right] - \prod_{i \leq k \leq i'} \left( 2^{-(k+1)+\varepsilon_n} \right)^{a_k} \rightarrow 0.$$

This immediately yields Theorem 1.1.

By the linearity of expectation, proving (2) reduces to understanding the probabilities

$$(3) \quad \mathbf{P}(\deg_{T_n}(v_k) = \lfloor \log n \rfloor + i_k, k \in [K])$$

for all  $i_1, \dots, i_K \in \mathbb{N}$  and  $v_1, \dots, v_K \in [n]$ ,  $K \in \mathbb{N}$ ; see Section 5 for more details.

In the standard model for RRT's described at the beginning,  $\deg_{T_n}(v)$  is a sum of Bernoulli variables:

$$\deg_{T_n}(v) = \sum_{v < u \leq n} \mathbf{1}_{\{u \rightarrow v\}}.$$

The lack of symmetry of the degrees  $\{\deg_{T_n}(v)\}_{v \in [n]}$  complicates the analysis of (3). In proving that  $\Delta_n / \log n \xrightarrow{\text{a.s.}} 1$ , Devroye and Lu [6] used that  $\{\deg_{T_n}(v)\}_{v \in [n]}$  are negatively orthant dependent (see [10] for a definition), which in particular means that for all  $S \subset [n]$  and  $m_1, \dots, m_n \in \mathbb{N}$

$$(4) \quad \mathbf{P}(\deg_{T_n}(v) \geq m_v, v \in S) \leq \prod_{v \in S} \mathbf{P}(\deg_{T_n}(v) \geq m_v)$$

and then obtained upper bounds for  $\mathbf{P}(\deg_{T_n}(v) \geq c \ln n)$  for each  $v \in [n]$ .

One approach to studying high degrees in  $T_n$  would be to obtain matching lower bounds for  $\mathbf{P}(\deg_{T_n}(v) \geq m_v, v \in S)$ , with uniform error terms even when  $m_v$  is large. Instead, we study trees  $T^{(n)}$ , mentioned in (1), above, for which we can obtain precise asymptotics for the analogous probabilities

$$(5) \quad \mathbf{P}(\deg_{T^{(n)}}(v) \geq m_v, v \in [K]).$$

The core of the paper lies in Proposition 4.2, which gives precise estimates of (5) for  $m_1, \dots, m_K$  in a suitable range. Broadly speaking,  $\deg_{T^{(n)}}(v)$  depends on a set of random *selection times*  $\mathcal{S}_v$  and the first streak of heads in a sequence of  $|\mathcal{S}_v|$  fair coin flips. As mentioned in the previous section, the degrees of  $T^{(n)}$  have the same distribution as the degrees in  $T_n$ . Consequently, our estimation of (5) allows us to obtain the following moments estimate.

**Proposition 2.1.** *For all  $c \in (0, 2)$  and  $K \in \mathbb{N}$  there is  $\alpha = \alpha(c, K) > 0$  such that the following holds. Fix any integers  $i, i'$  with  $0 < i + \log_n < i' + \log_n < c \ln n$ . Then for any non-negative integers  $a_i, \dots, a_{i'}$  with  $a_i + \dots + a_{i'} = K$ , we have*

$$\mathbf{E} \left[ (X_{\geq i'}^{(n)})_{a_{i'}} \prod_{i \leq k < i'} (X_k^{(n)})_{a_k} \right] = \left( 2^{-i'+\varepsilon_n} \right)^{a_{i'}} \prod_{i \leq k < i'} \left( 2^{-(k+1)+\varepsilon_n} \right)^{a_k} (1 + o(n^{-\alpha})).$$

Equipped with Proposition 2.1, the proofs of the theorems are straightforward. The rest of the paper is organized as follows. In Section 3, we explain how to define the trees  $T^{(n)}$  using Kingman's coalescent and establish the distributional relation between  $T^{(n)}$  and the RRT; see Corollary 3.4. In Section 4, we define the random sets  $(\mathcal{S}_v, v \in T^{(n)})$  and explain their relation with degrees in  $T^{(n)}$ . The proof of Proposition 4.2, which is our estimate of

(5), is then presented using a decoupling of the events in (5) and the concentration of the random variables  $|\mathcal{S}_v|$ . Finally, the proof of Proposition 2.1 is given in Section 5 and the proof of Theorems 1.2-1.4 are in Section 6.

### 3. RANDOM RECURSIVE TREES AND KINGMAN'S COALESCENT

In this section we give a representation of Kingman's coalescent in terms of labelled forests, and relate it to RRT's. All trees in the remainder of the paper are rooted, and we write  $r(t)$  for the root of tree  $t$ . By convention, edges of a tree are directed towards the root of the tree and we write  $uv$  to denote an edge directed from  $u$  to  $v$ . A forest  $f$  is a set of trees whose vertex sets are pairwise disjoint. The vertex set of a forest, denoted  $V(f)$ , is the union of the vertex sets of its trees. Similarly,  $E(f)$  denotes the set of edges in the trees of  $f$ . For  $n \geq 1$ , let

$$\mathcal{F}_n = \{f : V(f) = [n]\}$$

be the set of forests with vertex set  $[n]$ .

A sequence  $C = (f_1, \dots, f_n)$  of elements of  $\mathcal{F}_n$  is an  $n$ -chain if  $f_1$  is the forest in  $\mathcal{F}_n$  with  $n$  one-vertex trees and, for  $1 \leq i < n$ ,  $f_{i+1}$  is obtained from  $f_i$  by adding a directed edge between the roots of some pair of trees in  $f_i$ . If  $(f_1, \dots, f_n)$  is an  $n$ -chain then for  $1 \leq i \leq n$ , the forest  $f_i$  consists of  $n + 1 - i$  trees, and in this case we list its elements in increasing order of their smallest-labelled vertex as  $t_1^{(i)}, \dots, t_{n+1-i}^{(i)}$ .

**Definition 3.1.** *Kingman's  $n$ -coalescent is the random  $n$ -chain  $\mathbf{C} = (F_1, \dots, F_n)$  built as follows. Independently for each  $1 \leq i \leq n - 1$  let  $\{a_i, b_i\}$  be a random pair uniformly chosen from  $\{\{a, b\} : 1 \leq a < b \leq n + 1 - i\}$  and let  $\xi_i$  be independent with Bernoulli(1/2) distribution.*

*For  $1 \leq i < n$ , construct  $F_{i+1}$  from  $F_i$  as follows. If  $\xi_i = 1$  then add an edge from  $r(T_{b_i}^{(i)})$  to  $r(T_{a_i}^{(i)})$  and if  $\xi_i = 0$  then add an edge from  $r(T_{a_i}^{(i)})$  to  $r(T_{b_i}^{(i)})$ . The forest  $F_{i+1}$  consists of the new tree and the remaining  $n - 1 - i$  unaltered trees from  $F_i$ .*

For an example of the process see Figure 1.

**Lemma 3.2.** *Let  $\mathcal{CF}_n$  be the set of  $n$ -chains of elements in  $\mathcal{F}_n$ . Then  $|\mathcal{CF}_n| = n!(n - 1)!$  and Kingman's  $n$ -coalescent is a uniformly random element of  $\mathcal{CF}_n$ .*

*Proof.* Fix an  $n$ -chain  $(f_1, \dots, f_n) \in \mathcal{CF}_n$ . Then

$$\mathbf{P}((F_1, \dots, F_n) = (f_1, \dots, f_n)) = \prod_{k=1}^{n-1} \mathbf{P}(F_{k+1} = f_{k+1} | F_j = f_j, 1 \leq j \leq k).$$

Among the  $(n + 1 - k)(n - k)$  possible oriented edges between roots of  $f_k$ , there is exactly one whose addition yields  $f_{k+1}$ . It follows that the  $k$ -th term in the above product is  $((n + 1 - k)(n - k))^{-1}$ , so  $\mathbf{P}((F_1, \dots, F_n) = (f_1, \dots, f_n)) = [n!(n - 1)!]^{-1}$ . The result follows since this expression does not depend on  $(f_1, \dots, f_n) \in \mathcal{CF}_n$ .  $\square$

Recall that  $\mathcal{I}_n$  is the set of increasing trees with vertex set  $[n]$ . It is not difficult to see that  $|\mathcal{I}_n| = (n - 1)!$  and that a RRT is a uniformly random element of  $\mathcal{I}_n$ .

There is a natural mapping  $\phi$  between  $n$ -chains and increasing trees. Given an  $n$ -chain  $C = (f_1, \dots, f_n)$ , write  $t^{(n)} := t_1^{(n)}$  for the unique tree in  $f_n$ . Let  $L_C^- : E(t^{(n)}) \rightarrow [n - 1]$  be defined as follows. For each  $e \in E(t^{(n)})$ , let

$$L_C^-(e) = \max\{i \in [n - 1] : e \notin E(t^{(i)})\}.$$

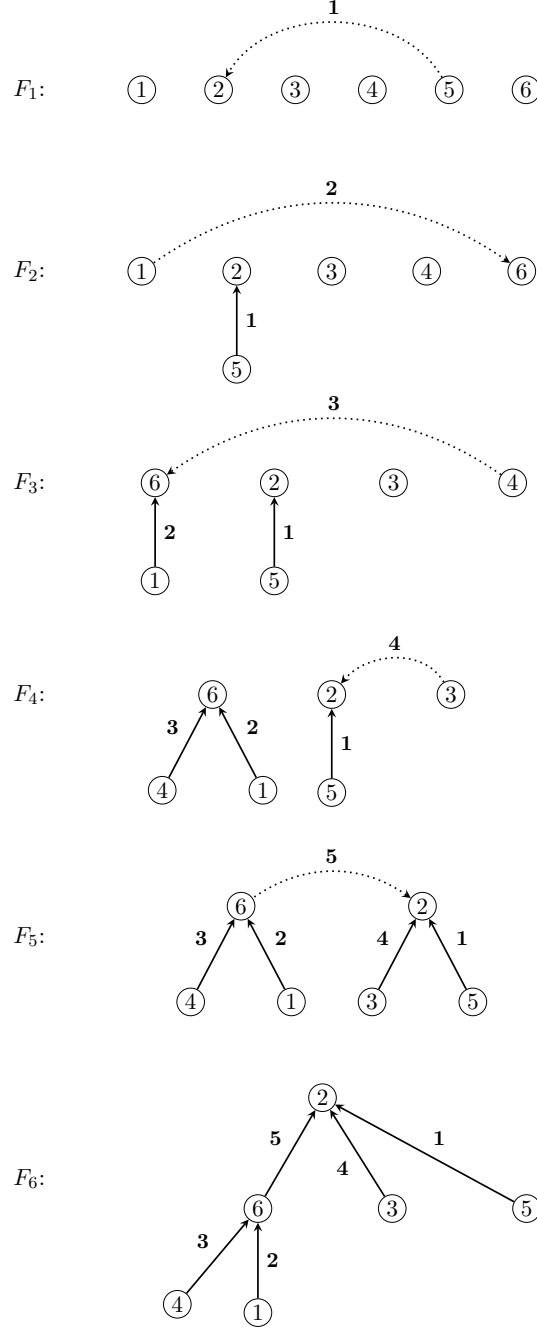


FIGURE 1. An example of Kingman's  $n$ -coalescent  $\mathbf{C} = (F_1, \dots, F_n)$  for  $n = 6$ . For  $1 \leq i < n$ ,  $F_i$  has, in dotted line, the edge in  $E(F_{i+1}) \setminus E(F_i)$ . Edges are marked with their time of addition; this is the function  $L_{\mathbf{C}}^-$  defined after Lemma 3.2. In this instance,  $\xi_1 = \xi_3 = \xi_4 = 1$ ,  $\xi_2 = \xi_5 = 0$  and  $\{a_1, b_1\} = \{2, 5\}$ ,  $\{a_2, b_2\} = \{1, 5\}$ ,  $\{a_3, b_3\} = \{1, 4\}$ ,  $\{a_4, b_4\} = \{2, 3\}$ ,  $\{a_5, b_5\} = \{1, 2\}$ .

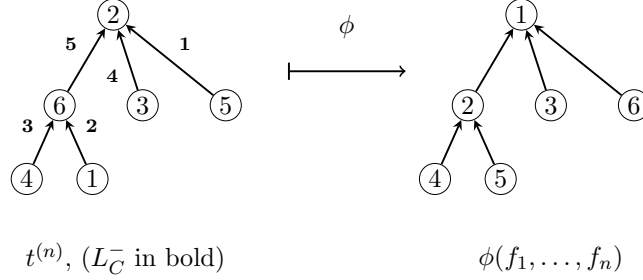


FIGURE 2. On the left a tree  $t^{(n)}$ ; edges are marked with  $L_C^-$ , from which the  $n$ -chain  $C = (f_1, \dots, f_n)$  can be recovered. On the right, the increasing tree  $\phi(f_1, \dots, f_n)$ ; it has the shape of  $t^{(n)}$  and the vertex labels  $\{L_C(v), v \in V(t^{(n)})\}$ .

We think of  $L_C^-$  as a function that keeps track of the *time of addition* of the edges along the  $n$ -chain  $C$ . Now, we define a vertex labelling  $L_C : V(t^{(n)}) \rightarrow [n]$  as follows. Let  $L_C(r(t^{(n)})) = 1$  and for each  $uv \in E(t^{(n)})$ , let

$$L_C(u) = n + 1 - L_C^-(uv);$$

then  $L_C(u)$  is the number of trees in the forest just before  $uv$  is added.

Note that for each  $i \in [n - 1]$ , the new edge in  $f_{i+1}$  joins the roots of two trees in  $f_i$  and is directed towards the root of the resulting tree. Thus, the labels  $\{L_C^-(e), e \in E(t^{(n)})\}$  increase along all paths in  $t^{(n)}$  towards the root  $r(t^{(n)})$  and consequently, the labels  $\{L_C(v), v \in V(t^{(n)})\}$  increase along root-to-leaf paths in  $t^{(n)}$ . This shows that relabelling the vertices of  $t^{(n)}$  with  $L_C$  yields an increasing tree (specifically, an element of  $\mathcal{I}_n$ ). See Figure 2 for an example.

**Proposition 3.3.** *Let  $\phi : \mathcal{CF} \rightarrow \mathcal{I}_n$  be defined as follows. For an  $n$ -chain  $C = (f_1, \dots, f_n)$  let  $\phi(C)$  be the tree obtained from  $t^{(n)}$  by relabelling its vertices with  $L_C$ . Then  $\phi(\mathbf{C})$ , the push-forward of Kingman's  $n$ -coalescent by  $\phi$ , has the law of a RRT of size  $n$ .*

*Proof.* First, we prove that  $\phi$  is onto. Fix an increasing tree  $t \in \mathcal{I}_n$ . For each  $j \in V(t) \setminus \{1\}$ , let  $v_j \in V(t)$  be such that  $ju_j \in E(t)$ , recall that edges are directed toward the root of  $t$ , thus  $v_j$  is uniquely defined. For each  $1 < j \leq n$ , let  $e_{n-j+1} = ju_j$ .

Now construct an  $n$ -chain  $C$  as follows. Let  $f_1$  be the forest with  $n$  one-vertex trees. For each  $1 < i \leq n$  construct  $f_i$  from  $f_{i-1}$  by adding the edge  $e_{i-1}$ . In other words, for each  $1 \leq i < n$ ,  $L_C^-(e_i) = i$  and so  $L_C(n + 1 - i) = n + 1 - L_C^-(e_i) = n + 1 - i$ ; also since  $r(t) = 1$ , we have  $L_C(1) = 1$ . Consequently,  $\phi(C) = t$ .

We claim that  $|\phi^{-1}(t)| \geq n!$  for any  $t \in \mathcal{I}_n$ . To see this, consider an  $n$ -chain  $C$  and a permutation  $\sigma : [n] \rightarrow [n]$ . Let  $C_\sigma$  be the  $n$ -chain obtained from  $C$  by permuting the vertices in each forest of  $C$  by  $\sigma$ . Since  $L_C(v)$  depends only on the time of addition of its outgoing edge (if any), it follows that  $\phi(C) = \phi(C_\sigma)$  for all permutations  $\sigma$ . By Lemma 3.2, this shows that  $\phi$  is  $n!$ -to-1 and that  $\phi(\mathbf{C})$  is a uniform element in  $\mathcal{I}_n$ .  $\square$

Since  $\phi(\mathbf{C})$  preserves the shape of  $T^{(n)}$  and only relabels its vertices, the degrees in  $T^{(n)}$  and  $\phi(\mathbf{C})$  are equal as multisets:  $\{\deg_{T^{(n)}}(v)\}_{v \in [n]} = \{\deg_{\phi(\mathbf{C})}(v)\}_{v \in [n]}$ . This immediately gives the following key corollary of Proposition 3.3, on which the rest of the paper relies.

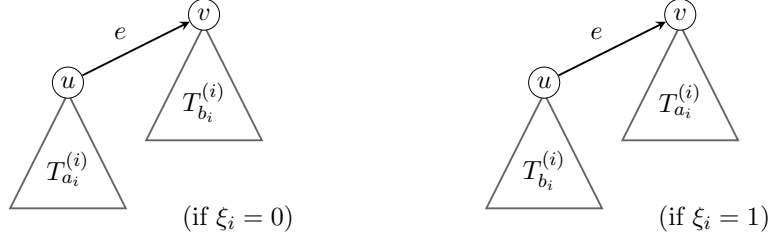


FIGURE 3. If  $v$  is a root in  $T_{a_i}^{(i)} \cup T_{b_i}^{(i)}$  and  $\xi_i$  favours  $v$ , then  $v$  increases its degree and remains a root in  $F_{i+1}$ .

**Corollary 3.4.** *For all  $n \in \mathbb{N}$ , we have the following equality in distribution holds jointly for all  $i \in \mathbb{Z}$ ,*

$$X_i^{(n)} \stackrel{d}{=} |\{v \in [n] : \deg_{T^{(n)}}(v) = \lfloor \log n \rfloor + i\}|.$$

We now proceed to the study of the joint distribution of the vertex degrees in  $T^{(n)}$ .

#### 4. DEGREE DISTRIBUTION: SELECTION SETS AND COIN FLIPS

By construction, the vertex degrees  $\{\deg_{T^{(n)}}(v)\}_{v \in [n]}$  are exchangeable. Our next goal is to explain how to approximate (5); that is, for any fixed  $k \in \mathbb{N}$  and integers  $m_1, \dots, m_k < 2 \ln n$ , to obtain estimates for  $\mathbf{P}(\deg_{T^{(n)}}(v) \geq m_v, v \in [k])$ .

The key to analyse the degrees in  $T^{(n)}$  is to understand how the degrees of a vertex  $v \in [n]$  change in Kingman's coalescent  $\mathbf{C} = (F_1, \dots, F_n)$ . For any vertex  $v$  and  $1 \leq i \leq n$ , denote  $\deg_{F_i}(v)$  the number of children of  $v$  in  $F_i$ . Also, we will simply write  $\deg(v) = \deg_{F_n}(v) = \deg_{T^{(n)}}(v)$ . For each  $1 \leq i < n$ , if  $\xi_i = 1$  we say that  $\xi_i$  *favours* the vertices of  $T_{a_i}^{(i)}$ , and otherwise that it favours the vertices of  $T_{b_i}^{(i)}$ . For  $v \in [n]$ , let

$$\mathcal{S}_v = \{i \in [n-1] : v \in T_{a_i}^{(i)} \cup T_{b_i}^{(i)}\}.$$

For any vertex  $v$ , and  $1 \leq i < n$ ,  $\deg_{F_{i+1}}(v)$  increases by one only if  $v$  is a root in  $F_i$ ,  $i \in \mathcal{S}_v$  and  $\xi_i$  favours  $v$ ; see Figure 3. Conversely, let  $p_v = \min\{i \in \mathcal{S}_v, \xi_i \text{ does not favour } v\}$ , then the first  $F_{i+1}$  in which  $v$  is not a root is exactly  $i = p_v$ . In this case, in  $F_{p_v+1}$  there is an outgoing edge from  $v$ , and  $v$  is not a root of any subsequent forests. As a consequence,  $\deg_{F_j}(v) = \deg_{F_{p_v}}(v)$  for  $p_v < j \leq n$ .

**Fact 4.1.** *For  $v \in [n]$ ,  $\deg(v) = \deg_{F_{p_v}}(v) = |\mathcal{S}_v \cap [p_v - 1]|$ .*

In other words,  $\deg(v)$  depends only on its first streak of favourable random variables  $\xi_i$  with  $i \in \mathcal{S}_v$ . More precisely, given  $|\mathcal{S}_v|$ , the degree  $\deg(v)$  is distributed as  $\min\{|\mathcal{S}_v|, G\}$ , where  $G$  is a Geometric(1/2) r.v. independent of  $\mathcal{S}_v$ .

Thus, it is relevant to observe that  $|\mathcal{S}_v|$  is distributed as an sum of independent (though not identically distributed) Bernoulli random variables and so it is concentrated around its mean  $\mathbf{E}[|\mathcal{S}_v|] = 2 \ln n + O(1)$ ; a more precise statement can be found in Proposition 4.5 below. Since  $|\mathcal{S}_v| \rightarrow \infty$  in probability as  $n \rightarrow \infty$ , it follows easily that  $\deg(v)$  is asymptotically geometric for any fixed node  $v$ . More strongly, the following proposition shows that for any fixed  $k$ , the random variables  $\{\deg_{T^{(n)}}(v)\}_{v \in [k]}$  asymptotically behave like independent Geometric random variables, even if they are conditioned to be quite large.



**Proposition 4.2.** *Fix  $c \in (0, 2)$  and  $k \in \mathbb{N}$ . There exists  $\alpha = \alpha(c, k) > 0$  such that uniformly over positive integers  $m_1, \dots, m_k < c \ln n$ ,*

$$\mathbf{P}(\deg_{T^{(n)}}(v) \geq m_v, v \in [k]) = 2^{-\sum_v m_v} (1 + o(n^{-\alpha})).$$

We now explain how the events in the proposition above can be decoupled into a product of two probabilities, one of them corresponding to tail bounds for the random variables  $|\mathcal{S}_v|$ . We start with an upper bound for Proposition 4.2.

**Lemma 4.3.** *For any  $k \in \mathbb{N}$  and positive integers  $m_1, \dots, m_k < n$ ,*

$$\mathbf{P}(\deg(v) \geq m_v, v \in [k]) \leq 2^{-\sum_v m_v} \mathbf{P}(|\mathcal{S}_v| \geq m_v, v \in [k]).$$

*Equality holds for  $k = 1$ .*

*Proof.* For each  $v \in [k]$  list  $\mathcal{S}_v$  in increasing order as  $(i_{v,j}, 1 \leq j \leq |\mathcal{S}_v|)$ . Let  $\mathcal{A}$  be the set of sequences  $A = (A_1, \dots, A_k)$  satisfying  $A_v \subset [n-1]$  and  $|A_v| = m_v$  for all  $v \in [k]$ . For every  $A \in \mathcal{A}$ , let  $D_A$  be the event that  $|\mathcal{S}_v| \geq m_v$  and  $\{i_{v,1}, \dots, i_{v,m_v}\} = A_v$ , for all  $v \in [k]$ . By Fact 4.1, if  $\deg(v) \geq m_v$  then necessarily  $|\mathcal{S}_v| \geq m_v$  so

$$\{\deg(v) \geq m_v, v \in [k]\} \cap D_A = \{\xi_{i_{v,j}} \text{ favours } v \text{ for all } j \in [m_v], v \in [k]\} \cap D_A.$$

Now,  $\xi_i$  are i.i.d Bernoulli(1/2) r.v.'s. Thus, if  $D_A$  has positive probability then

$$\mathbf{P}(\xi_{i_{v,j}} \text{ favours } v \text{ for all } j \in [m_v], v \in [k] | D_A) = \begin{cases} 2^{-\sum_v m_v} & \text{if } |A_u \cap A_v| = 0, \forall u \neq v \in [k] \\ 0 & \text{o.w.} \end{cases}$$

The second case follows from the fact that if  $i \in \mathcal{S}_u \cap \mathcal{S}_v$  for some  $u \neq v$ , then  $\xi_i$  cannot favour both  $u$  and  $v$ . The events  $(D_A, A \in \mathcal{A})$  are pairwise disjoint, and if  $\deg(v) \geq m_v$  for all  $v \in [k]$  then one of the events  $D_A$  must occur. It follows that

$$\begin{aligned} \mathbf{P}(\deg(v) \geq m_v, v \in [k]) &= \sum_{A \in \mathcal{A}} \mathbf{P}(D_A, \deg(v) \geq m_v, v \in [k]) \\ &\leq \sum_{A \in \mathcal{A}} 2^{-\sum_v m_v} \mathbf{P}(D_A) \\ &= 2^{-\sum_v m_v} \mathbf{P}(|\mathcal{S}_v| \geq m_v, v \in [k]). \end{aligned}$$

Finally, the second line holds with equality when  $k = 1$ .  $\square$

For the lower bound we restrict to events  $D_A$  where the sets  $A_v$  are already disjoint. To do so, we consider instead the vertex degrees in  $F_I$  for some  $I < n$ . For  $k \geq 2$  let

$$\tau_k = \min\{i \in [n-1] : \{a_i, b_i\} \subset [k]\}.$$

Since  $F_i \subset F_j$  for all  $i \leq j \in [n]$  we have that for any  $I < n$

$$\begin{aligned} \mathbf{P}(\deg(v) \geq m_v, v \in [k]) &\geq \mathbf{P}(\deg_{F_{I+1}}(v) \geq m_v, v \in [k]) \\ (6) \quad &\geq \mathbf{P}(I < \tau_k, \deg_{F_{I+1}}(v) \geq m_v, v \in [k]). \end{aligned}$$

Recall that trees in  $F_i$  are listed in increasing order of their least elements; this implies that indices of the trees of vertices  $1, \dots, k$  do not change until two trees indexed by  $a, b \leq k$  are merged. Therefore, for all  $v \in [k]$ ,  $v \in T_v^{(i)}$  for  $i \leq \tau_k$ . This implies the sets  $\{\mathcal{S}_v \cap [\tau_k - 1], v \in [k]\}$  are pairwise disjoint. These observations allow us to obtain a lower bound analogous to Lemma 4.3.

**Lemma 4.4.** *For any positive integers  $k \geq 2$  and  $m_1, \dots, m_k, I < n$ ,*

$$\mathbf{P}(\deg(v) \geq m_v, v \in [k]) \geq 2^{-\sum_v m_v} \mathbf{P}(I < \tau_k, |S_v \cap [I]| \geq m_v, v \in [k]).$$

*Proof.* By (6), it suffices to bound  $\mathbf{P}(I < \tau_k, \deg_{F_{I+1}}(v) \geq m_v, v \in [k])$ .

Let  $\mathcal{A}^*$  be the set of sequences  $A = (A_1, \dots, A_k)$  of pairwise disjoint subsets of  $[I]$  satisfying  $|A_v| = m_v$  for all  $v \in [k]$ . For each  $A \in \mathcal{A}^*$ , let  $D_A$  be the event that for all  $v \in [k]$ ,  $\{i_{v,j}, \dots, i_{v,m_v}\} = A_v$  (and so  $|S_v \cap [I]| \geq m_v$ ).

As in the proof of Lemma 4.3, we have that

$$\{\deg_{F_{I+1}}(v) \geq m_v, v \in [k]\} \cap D_A = \{\xi_{i_{v,j}} \text{ favours } v \text{ for all } j \in [m_v], v \in [k]\} \cap D_A.$$

In this case, the sets  $A_v$  are pairwise disjoint. If  $\mathbf{P}(D_A) > 0$  then

$$\mathbf{P}(\xi_{i_{v,j}} \text{ favours } v \text{ for all } j \in [m_v], v \in [k] | D_A) = 2^{-\sum_v m_v}.$$

Recall that  $I < \tau_k$  if and only if the sets  $\{S_v \cap [I], v \in [k]\}$  are pairwise disjoint; that is, if one of the events  $D_A$  occur. We then have

$$\begin{aligned} \mathbf{P}(I < \tau_k, \deg_{F_{I+1}}(v) \geq m_v, v \in [k]) &= \sum_{A \in \mathcal{A}^*} \mathbf{P}(D_A, \deg_{F_{I+1}}(v) \geq m_v, v \in [k]) \\ &= \sum_{A \in \mathcal{A}^*} 2^{-\sum_v m_v} \mathbf{P}(D_A) \\ &= 2^{-\sum_v m_v} \mathbf{P}(I < \tau_k, |S_v \cap [I]| \geq m_v, v \in [k]). \quad \square \end{aligned}$$

To use Lemma 4.4 we need tail bounds for  $|S_v \cap [I]|$  for some suitable  $I < n$ ; these are provided by the following proposition.

**Proposition 4.5.** *Fix  $\varepsilon \in (0, 1)$  and  $c \in (0, 2(1 - \varepsilon))$ . Then there exists  $\beta = \beta(c, \varepsilon) > 0$  such that for any vertex  $v$ ,*

$$\mathbf{P}(|S_v \cap [n - \lceil n^\varepsilon \rceil]| < c \ln n) = o(n^{-\beta}).$$

*Proof.* Fix  $\varepsilon \in (0, 1)$  and  $c \in (0, 2(1 - \varepsilon))$ . Let  $\{B_i, i \in \mathbb{N}\}$  be a collection of independent Bernoulli r.v.'s, with  $\mathbf{E}[B_i] = \frac{2}{i}$ . Recall the definition of  $S_v$  at the beginning of the section.

For any fixed vertex  $v \in [n]$ , and each  $i \in [n - 1]$ , the probability of the event  $\{v \in T_{a_i}^{(i)} \cup T_{b_i}^{(i)}\}$  is  $2/(n - i + 1)$ ; this is because, in the forest  $F_i$ , there are  $n - i + 1$  trees and the trees  $T_{a_i}^{(i)}, T_{b_i}^{(i)}$  are chosen uniformly at random among them. Since each of these events are independent we have  $|S_v| \stackrel{d}{=} \sum_{i=2}^n B_i$ . Moreover, writing  $W_{n,\varepsilon} = \sum_{i=n-\lceil n^\varepsilon \rceil}^n B_i$ , we also have

$$W_{n,\varepsilon} \stackrel{d}{=} |S_v \cap [n - \lceil n^\varepsilon \rceil]|.$$

We now apply Bernstein's inequality (see, e.g., [9], Theorem 2.8) to obtain that for any  $t > 0$ ,

$$\mathbf{P}(W_{n,\varepsilon} \leq \mathbf{E}[W_{n,\varepsilon}] - t) \leq \exp \left\{ -\frac{t^2}{2\mathbf{E}[W_{n,\varepsilon}]} \right\}.$$

We take  $t = \mathbf{E}[W_{n,\varepsilon}] - c \ln n$ . Since

$$\mathbf{E}[W_{n,\varepsilon}] = \sum_{i=n-\lceil n^\varepsilon \rceil}^n \frac{2}{i} = 2(1 - \varepsilon) \ln n + O(1),$$

setting  $\delta = 2(1 - \varepsilon) - c > 0$  we have  $t = \delta \ln n + O(1)$ , so

$$\mathbf{P}(|S_v \cap [n - \lceil n^\varepsilon \rceil]| < c \ln n) = \mathbf{P}(W_{n,\varepsilon} \leq \mathbf{E}[W_{n,\varepsilon}] - t) = O(1) \cdot n^{-\delta^2/(4(1-\varepsilon))}.$$

Choosing  $0 < \beta < \delta^2/4(1 - \varepsilon)$ , the result follows.  $\square$

The following lemma is the last ingredient for Proposition 4.2.

**Lemma 4.6.** *Fix an integer  $k \geq 2$  and let  $\varepsilon \in (0, 1)$ . Then, for  $n$  large enough,*

$$\mathbf{P}(\tau_k \leq n - \lceil n^\varepsilon \rceil) \leq \frac{2k^2}{\lceil n^\varepsilon \rceil - 1}.$$

*Proof.* By the definition of  $\tau_k$ , if  $\tau_k > n - \lceil n^\varepsilon \rceil$  then  $\{a_i, b_i\} \not\subset [k]$  for all  $1 \leq i \leq n - \lceil n^\varepsilon \rceil$ . The events that  $\{a_i, b_i\} \not\subset [k]$  are independent for distinct  $i$  and  $\mathbf{P}(\{a_i, b_i\} \subset [k]) = \frac{k(k-1)}{(n+1-i)(n-i)}$ , so we have that

$$\mathbf{P}(\tau_k > n - \lceil n^\varepsilon \rceil) = \prod_{i=1}^{n - \lceil n^\varepsilon \rceil} \left(1 - \frac{k(k-1)}{(n+1-i)(n-i)}\right) \geq 1 - \sum_{i=1}^{n - \lceil n^\varepsilon \rceil} \frac{2k^2}{(n-i)^2}$$

The last inequality holds for  $n$  large enough. Since  $\sum_{j=m}^{\infty} j^{-2} \leq \int_{m-1}^{\infty} x^{-2} dx = (m-1)^{-1}$ , we get

$$\mathbf{P}(\tau_k \leq n - \lceil n^\varepsilon \rceil) \leq \sum_{i=1}^{n - \lceil n^\varepsilon \rceil} \frac{2k^2}{(n-i)^2} \leq \sum_{j=\lceil n^\varepsilon \rceil}^{\infty} \frac{2k^2}{j^2} = \frac{2k^2}{\lceil n^\varepsilon \rceil - 1}. \quad \square$$

We finish this section with the proof of Proposition 4.2.

*Proof of Proposition 4.2.* Fix  $c \in (0, 2)$ ,  $k \in \mathbb{N}$  and let  $m_1, \dots, m_k < c \ln n$  be positive integers. Let  $\varepsilon = (2 - c)/4$  so that Proposition 4.5 holds for some  $\beta(c) = \beta(c, \varepsilon) > 0$ . For  $k = 1$ , the result follows from the equality in Lemma 4.3 and Proposition 4.5 since

$$\mathbf{P}(|\mathcal{S}_1| < m_1) \leq \mathbf{P}(|\mathcal{S}_1 \cap [n - \lceil n^\varepsilon \rceil]| < c \ln n) = o(n^{-\beta}).$$

For  $k \geq 2$ , the upper bound is likewise established immediately by Lemma 4.3. For the lower bound, letting  $I = n - \lceil n^\varepsilon \rceil$ , by Lemma 4.6 and Proposition 4.5 we have

$$\mathbf{P}(I < \tau_k, |\mathcal{S}_v \cap [I]| \geq m_v, v \in [k]) \geq 1 - \mathbf{P}(I \geq \tau_k) - \sum_{v \in [k]} \mathbf{P}(|\mathcal{S}_v \cap [I]| < m_v) \geq 1 - o(n^{-\alpha}),$$

where  $\alpha < \min\{\beta, \varepsilon\}$ . By Lemma 4.4, it follows that

$$\mathbf{P}(\deg(v) \geq m_v, v \in [k]) = 2^{-\sum_v m_v} (1 + o(n^{-\alpha})),$$

as required.  $\square$

## 5. PROOF OF PROPOSITION 2.1

By Corollary 3.4 we can study vertex degrees in  $T^{(n)}$  and derive conclusions about the variables  $X_i^{(n)}, X_{\geq i}^{(n)}, i \in \mathbb{Z}$ . Recall that we write  $\deg(v) = \deg_{T^{(n)}}(v)$ , for  $v \in [n]$ .

**Lemma 5.1.** *For any  $k \in \mathbb{N}$  and integers  $m_1, \dots, m_k$ ,*

$$\mathbf{P}(\deg(u) = m_u, u \in [k]) = \sum_{j=0}^k \sum_{\substack{S \subset [k] \\ |S|=j}} (-1)^j \mathbf{P}(\deg(u) \geq m_u + \mathbf{1}_{[u \in S]}, u \in [k]).$$

Furthermore, for  $k' \in \mathbb{N}$  and integers  $m_{k+1}, \dots, m_{k+k'}$ ,

$$\begin{aligned} & \mathbf{P}(\deg(u) = m_u, \deg(v) \geq m_v, 1 \leq u \leq k < v \leq k + k') \\ &= \sum_{j=0}^k \sum_{\substack{S \subseteq [k] \\ |S|=j}} (-1)^j \mathbf{P}(\deg(v) \geq m_v + \mathbf{1}_{[v \in S]}, v \in [k + k']). \end{aligned}$$

*Proof.* The second equation follows by intersecting the event  $\{\deg(v) \geq m_v, k < v \leq k + k'\}$  along all probabilities in the first equation. The first is straightforwardly proved using the inclusion-exclusion principle.  $\square$

We are now ready to prove Proposition 2.1.

*Proof of Proposition 2.1.* Let  $c \in (0, 2)$  and  $K \in \mathbb{N}$ . Let  $i < i'$  be integers such that  $0 < i + \log_n < i' + \log_n < c \ln n$  and let  $a_j$ ,  $i \leq j \leq i'$  be non-negative integers with  $a_i + \dots + a_{i'} = K$ . We are interested in the factorial moments  $\mathbf{E} \left[ (X_{\geq i'}^{(n)})_{a_{i'}} \prod_{i \leq k < i'} (X_k^{(n)})_{a_k} \right]$ .

For  $i \leq k \leq i'$ , for each  $v$  with  $\sum_{l=i}^{k-1} a_l < v \leq \sum_{l=i}^k a_l$  let  $m_v = \lfloor \log n \rfloor + k$ . Let  $K' = K - a_{i'}$ , by Corollary 3.4 and the exchangeability of the vertex degrees of  $T^{(n)}$ ,

$$\begin{aligned} \mathbf{E} \left[ (X_{\geq i'}^{(n)})_{a_{i'}} \prod_{i \leq k < i'} (X_k^{(n)})_{a_k} \right] &= (n)_K \mathbf{P}(\deg(u) = m_u, \deg(v) \geq m_v, 1 \leq u \leq K' < v \leq K) \\ &= (n)_K \sum_{l=0}^{K'} \sum_{\substack{S \subseteq [K'] \\ |S|=l}} (-1)^l \mathbf{P}(\deg(v) \geq m_v + \mathbf{1}_{[v \in S]}, v \in [K]), \end{aligned}$$

the last equality by Lemma 5.1. At this point we can apply Proposition 4.2 to each of the terms. Since  $m_v \leq c \ln n$  for  $v \in [K]$ , there is  $\alpha' = \alpha'(c, K) > 0$  such that

$$\begin{aligned} & \sum_{l=0}^{K'} \sum_{\substack{S \subseteq [K'] \\ |S|=l}} (-1)^l \mathbf{P}(\deg(v) \geq m_v + \mathbf{1}_{[v \in S]}, v \in [K]) \\ &= \sum_{l=0}^{K'} \sum_{\substack{S \subseteq [K'] \\ |S|=l}} (-1)^l 2^{-l - \sum_v m_v} (1 + o(n^{-\alpha'})) \\ &= 2^{-\sum_v m_v} (1 + o(n^{-\alpha'})) \sum_{l=0}^{K'} \sum_{\substack{S \subseteq [K'] \\ |S|=l}} (-1)^l 2^{-l} \\ &= 2^{-K' - \sum_v m_v} (1 + o(n^{-\alpha'})). \end{aligned}$$

Using that  $(n)_K = n^K (1 + o(n^{-1}))$ , we get

$$\mathbf{E} \left[ (X_{\geq i'}^{(n)})_{a_{i'}} \prod_{i \leq k < i'} (X_k^{(n)})_{a_k} \right] = 2^{K \log n - K' - \sum_{v=1}^K m_v} (1 + o(n^{-\alpha}));$$

where  $\alpha = \min\{\alpha', 1\}$ . Finally, to complete the proof, note that

$$\begin{aligned} K \log n - K' - \sum_{v=1}^K m_v &= \sum_{v=K'+1}^K (\log n - m_v) + \sum_{v=1}^{K'} (\log n - 1 - m_v) \\ &= (-i' + \varepsilon_n) a_{i'} + \sum_{k=i}^{i'-1} (-k - 1 + \varepsilon_n) a_k. \end{aligned} \quad \square$$

## 6. PROOFS OF THE MAIN THEOREMS

*Proof of Theorem 1.2.* By Theorem 11.1.VII of [4], weak convergence in  $\mathcal{M}_{\mathbb{Z}^*}^\#$  is equivalent to convergence of FDD's, that is, convergence of every finite family of bounded continuity sets; see Definition 11.1.IV of [4]. For any point process  $\xi$  on  $\mathbb{Z}$  and any  $i \in \mathbb{Z}$ , we have that  $\mathbb{Z} \cap [i, \infty)$  is a bounded stochastic continuity set for the underlying measure of  $\xi$  in  $\mathcal{M}_{\mathbb{Z}^*}^\#$ . Thus, any FDD of  $\xi$  can be recovered from suitable marginals of the joint distribution of  $(\xi(i), \dots, \xi(i-1'), \xi[i, \infty))$  for some  $i < i' \in \mathbb{Z}$ .

Let  $\varepsilon \in [0, 1]$  and  $(n_l)_{l \geq 1}$  be an increasing sequence with  $\varepsilon_{n_l} \rightarrow \varepsilon$ . The goal then is to prove that, for any integers  $i < i'$ , the joint distribution of

$$X_i^{(n_l)}, \dots, X_{i'-1}^{(n_l)}, X_{\geq i'}^{(n_l)}$$

converges to the joint distribution of

$$\mathcal{P}^\varepsilon(i), \dots, \mathcal{P}^\varepsilon(i' - 1), \mathcal{P}^\varepsilon[i', \infty),$$

that is, to the law of independent Poisson r.v.'s with parameters  $2^{-i-1+\varepsilon}, \dots, 2^{-i'-2+\varepsilon}, 2^{-i'+\varepsilon}$ .

We compute the limit of the factorial moments of  $X_i^{(n_l)}, \dots, X_{i'-1}^{(n_l)}, X_{\geq i'}^{(n_l)}$ . For any non-negative integers  $a_i, \dots, a_{i'}$ , by Proposition 2.1,

$$\begin{aligned} \mathbf{E} \left[ (X_{\geq i'}^{(n)})_{a_{i'}} \prod_{i \leq k < i'} (X_k^{(n)})_{a_k} \right] &= \left( 2^{-i'+\varepsilon_n} \right)^{a_{i'}} \prod_{i \leq k < i'} \left( 2^{-(k+1)+\varepsilon_n} \right)^{a_k} (1 + o(n^{-\alpha})) \\ &\rightarrow \left( 2^{-i'+\varepsilon} \right)^{a_{i'}} \prod_{i \leq k < i'} \left( 2^{-(k+1)+\varepsilon} \right)^{a_k}, \end{aligned}$$

as  $n_l \rightarrow \infty$ . The limit correspond to the factorial moment

$$\mathbf{E} \left[ (\mathcal{P}^\varepsilon[i', \infty))_{a_{i'}} \prod_{i \leq k < i'} (\mathcal{P}^\varepsilon(k))_{a_k} \right].$$

The result follows (by, e.g. Theorem 6.10 of [9]).  $\square$

*Proof of Theorem 1.3.* Since  $\{\Delta_n \geq \lfloor \log n \rfloor + i\} = \{X_{\geq i}^{(n)} > 0\}$ , we need only to estimate  $\mathbf{P}(X_{\geq i}^{(n)} > 0)$ . If  $i = O(1)$ , then  $\exp\{-2^{-i+\varepsilon_n}\} = O(1)$  and so it suffices to prove that

$$\mathbf{P}(X_{\geq i}^{(n)} = 0) - \exp\{-2^{-i+\varepsilon_n}\} \rightarrow 0,$$

as  $n \rightarrow \infty$ . This follows from Theorem 1.2 and the subsubsequence principle. Suppose that there exists  $\delta > 0$  and a subsequence  $n_k$  for which  $|\mathbf{P}(X_{\geq i}^{(n_k)} = 0) - \exp\{-2^{-i+\varepsilon_{n_k}}\}| > \delta$ . Since  $\{\varepsilon_{n_k}\}_{k \geq 1}$  is a bounded set there is a subsubsequence  $n_{k_l}$  such that  $\varepsilon_{n_{k_l}} \rightarrow \varepsilon$  for some  $\varepsilon \in [0, 1]$ . By Theorem 1.2,  $\mathbf{P}(X_{\geq i}^{(n_{k_l})} = 0) \rightarrow \exp\{-2^{-i+\varepsilon}\}$ ; this contradicts our assumption on the subsequence  $n_k$ .

Now consider the case  $i \rightarrow \infty$  with  $i + \log_n < 2 \ln n$ . By a standard inclusion-exclusion argument (see, e.g., [3] Corollary 1.11),

$$(7) \quad \mathbf{P} \left( X_{\geq i}^{(n)} = 0 \right) = \sum_{r=0}^n (-1)^r \frac{\mathbf{E} \left[ (X_{\geq i}^{(n)})_r \right]}{r!},$$

and this sum has the so called *alternating inequalities* property; this means that partial sums alternatively serve as upper and lower bounds for  $\mathbf{P} \left( X_{\geq i}^{(n)} = 0 \right)$ . Consequently <sup>1</sup>,

$$(8) \quad \mathbf{E} \left[ X_{\geq i}^{(n)} \right] - \frac{1}{2} \mathbf{E} \left[ (X_{\geq i}^{(n)})_2 \right] \leq \mathbf{P} \left( X_{\geq i}^{(n)} > 0 \right) \leq \mathbf{E} \left[ X_{\geq i}^{(n)} \right].$$

Using Proposition 2.1 and the fact that  $i \rightarrow \infty$ , we have that  $\mathbf{E} \left[ X_{\geq i}^{(n)} \right] = 2^{-i+\varepsilon_n} (1 + o(1))$  and

$$\mathbf{E} \left[ X_{\geq i}^{(n)} \right] - \frac{1}{2} \mathbf{E} \left[ (X_{\geq i}^{(n)})_2 \right] = 2^{-i+\varepsilon_n} (1 + o(1)) = (1 - \exp\{-2^{-i+\varepsilon_n}\})(1 + o(1)).$$

The result follows.  $\square$

*Proof of Theorem 1.4.* We again use the method of moments. By Theorem 1.24 of [3], it suffices to prove that, as  $n \rightarrow \infty$

$$(9) \quad \mathbf{E} \left[ (X_i^{(n)})_a \right] - (2^{-i-1+\varepsilon_n})^a = o(2^{-(i+1-\varepsilon_n)b}),$$

for all fixed  $1 \leq a \leq b$ . Since  $i = o(\ln n)$ , we have that  $2^{-i-1+\varepsilon_n} = n^{o(1)}$ . On the other hand, by Proposition 2.1 there is  $\alpha > 0$  such that

$$\mathbf{E} \left[ (X_i^{(n)})_a \right] - (2^{-i-1+\varepsilon_n})^a = o(n^{-\alpha} 2^{-(i+\varepsilon_n)a}) = n^{-\alpha+o(1)} = o(n^{o(1)}).$$

Therefore, condition (9) is satisfied and the proof is complete.  $\square$

**Acknowledgements.** Laura Eslava would like to thank Henning Sulzbach for some very helpful discussions.

## REFERENCES

- [1] Louigi Addario-Berry. Partition functions of discrete coalescents: from Cayley's formula to Frieze's  $\zeta(3)$  limit theorem. In *XI Symposium on Probability and Stochastic Processes*, volume 68 of *Progress in Probability*, Basel, 2015. Birkhauser.
- [2] Nathanaël Berestycki. *Recent progress in coalescent theory*, volume 16 of *Ensaïos Matemáticos [Mathematical Surveys]*. Sociedade Brasileira de Matemática, Rio de Janeiro, 2009. ISBN 978-85-85818-40-1.
- [3] Béla Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001. ISBN 0-521-80920-7; 0-521-79722-5. doi: 10.1017/CBO9780511814068. URL <http://dx.doi.org/10.1017/CBO9780511814068>.
- [4] D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. II. Probability and its Applications* (New York). Springer, New York, second edition, 2008. ISBN 978-0-387-21337-8. doi: 10.1007/978-0-387-49835-5. URL <http://dx.doi.org/10.1007/978-0-387-49835-5>. General theory and structure.

---

<sup>1</sup>A similar lower bound for  $\mathbf{P} \left( X_{\geq i}^{(n)} > 0 \right)$  could be obtained from Paley-Zigmond's inequality.

- [5] L. Devroye. Branching processes in the analysis of the heights of trees. *Acta Inform.*, 24(3):277–298, 1987. ISSN 0001-5903. doi: 10.1007/BF00265991. URL <http://dx.doi.org/10.1007/BF00265991>.
- [6] Luc Devroye and Jiang Lu. The strong convergence of maximal degrees in uniform random recursive trees and dags. *Random Structures Algorithms*, 7(1):1–14, 1995. ISSN 1042-9832. doi: 10.1002/rsa.3240070102. URL <http://dx.doi.org/10.1002/rsa.3240070102>.
- [7] William Goh and Eric Schmutz. Limit distribution for the maximum degree of a random recursive tree. *J. Comput. Appl. Math.*, 142(1):61–82, 2002. ISSN 0377-0427. doi: 10.1016/S0377-0427(01)00460-5. URL [http://dx.doi.org/10.1016/S0377-0427\(01\)00460-5](http://dx.doi.org/10.1016/S0377-0427(01)00460-5). Probabilistic methods in combinatorics and combinatorial optimization.
- [8] Svante Janson. Asymptotic degree distribution in random recursive trees. *Random Structures Algorithms*, 26(1-2):69–83, 2005. ISSN 1042-9832. doi: 10.1002/rsa.20046. URL <http://dx.doi.org/10.1002/rsa.20046>.
- [9] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000. ISBN 0-471-17541-2. doi: 10.1002/9781118032718. URL <http://dx.doi.org/10.1002/9781118032718>.
- [10] Kumar Joag-Dev and Frank Proschan. Negative association of random variables, with applications. *Ann. Statist.*, 11(1):286–295, 1983. ISSN 0090-5364. doi: 10.1214/aos/1176346079. URL <http://dx.doi.org/10.1214/aos/1176346079>.
- [11] Hosam M. Mahmoud and R. T. Smythe. Asymptotic joint normality of outdegrees of nodes in random recursive trees. *Random Structures Algorithms*, 3(3):255–266, 1992. ISSN 1042-9832. doi: 10.1002/rsa.3240030305. URL <http://dx.doi.org/10.1002/rsa.3240030305>.
- [12] A. Meir and J. W. Moon. Recursive trees with no nodes of out-degree one. *Congr. Numerantium* 66: 49–62, 1988.
- [13] Hwa Sung Na and Anatol Rapoport. Distribution of nodes of a tree by degree. *Math. Biosci.*, 6:313–329, 1970. ISSN 0025-5564.
- [14] Boris Pittel. Note on the heights of random recursive trees and random  $m$ -ary search trees. *Random Structures Algorithms*, 5(2):337–347, 1994. ISSN 1042-9832. doi: 10.1002/rsa.3240050207. URL <http://dx.doi.org/10.1002/rsa.3240050207>.
- [15] Jerzy Szymański. On the maximum degree and the height of a random recursive tree. In *Random graphs '87 (Poznań, 1987)*, pages 313–324. Wiley, Chichester, 1990.

*E-mail address:* [louigi@problab.ca](mailto:louigi@problab.ca)

*E-mail address:* [laura.eslavafernandez@mail.mcgill.ca](mailto:laura.eslavafernandez@mail.mcgill.ca)

DEPARTMENT OF MATHEMATICS AND STATISTICS, MCGILL UNIVERSITY, MONTREAL, CANADA